PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS
The Patent Office

INVESTOR IN PEOPLE

PCT/EP03/12830

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 3 November 2003

18NOV02 E763971-1 002029
P01/7700 0.00-0226722.7

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

1/77

THE PATENT OFFICE
D
15 NOV 2002
LONDON

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference — RJE/RSG/PG5029

2. Patent application number *(The Patent Office will fill in his part)* — 0226722.7 — 15 NOV 2002

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

   Glaxo Group Limited
   Glaxo Wellcome House, Berkeley Avenue,
   Greenford, Middlesex UB6 0NN, Great Britain

   Patents ADP number *(if you know it)*

   If the applicant is a corporate body, give the country/state of its incorporation — United Kingdom

   00435 87002

4. Title of the invention — Vaccine

5. Name of your agent *(if you have one)* — Corporate Intellectual Property

   "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

   GlaxoSmithKline
   Corporate Intellectual Property (CN9 25.1)
   980 Great West Road
   BRENTFORD
   Middlesex TW8 9GS

   Patents ADP number *(if you know it)* — 07960 982003

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right

grant of a patent required in support of this request? *(Answer yes if:*

a) any applicant named in part 3 is not an inventor, or

b) there is an inventor who is named as an applicant, or

c) any named applicant is a corporate body

See note (d)

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	0
Description	24
Claim(s)	~~1~~
Abstract	0
Drawings	17 only

10. If you are also filing any of the following, state how many against each item.

Priority Documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	Claims are included on last page of description

11. We request the grant of a patent on the basis of this application

Signature [signature] Date 15-Nov-02

R J Easeman

12. Name and daytime telephone number of person to contact in the United Kingdom: R J Easeman 020 80474407

Warning

After an application for a Patent has beeen filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission unless an application has been filed at least six weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505
b) Write your answers in capital letters using black ink or you may type them.
c) If there is not enough space for all relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be

attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

f) For details of the fee and ways to pay please contact the Patent Office.

Patents Form 1/77

Vaccine

The present invention relates to methods and compositions useful in the treatment and prevention of Hepatitis C virus (HCV) infections and the symptoms and diseases associated therewith. In particular the present invention relates to DNA vaccines comprising polynucleotide sequences encoding HCV proteins, and methods of treatment of individuals infected with HCV comprising administration of the vaccines of the present invention.

HCV was identified recently as the leading causative agent of post-transfusion and community acquired non A, non B hepatitis. Approximately 170m people are chronically infected with HCV, with prevalence between 1-10%. The health care cost in the US, where the prevalence is 1.8%, is estimated to be $2 billion. Between 40-60% of liver disease is due to HCV and 30% UK transplants are for HCV infections. Although HCV is initially a sub-clinical infection more than 90% of patients develop chronic disease. The disease process typically develops from chronic active hepatitis (70%), fibrosis, cirrhosis (40%) to hepato-cellular carcinoma (60%). Infection to cirrhosis has a median time of 20 years and that for hepato-cellular carcinoma of 20 years (Lauer G.and Walker B. 2001, Hepatits C virus Infection. N Engl J. Med 345, 41, Cohen J. 2001. The Scientific challenge of Hepatitis C. Science 285 (5424) 26.

There is a great need for the improved treatment of HCV. There are currently no small molecule replication inhibitors available. The current gold standard of ribovirin and PEGylated interferon represents the mainstay for treating HCV infection. However the ability of the current regimens to achieve sustained response remains sub-optimal (overall 50% response rate for up to 6 months, however, for genotype 1b the response rate is lower (27%). This treatment is also associated with unpleasant side effects. This results in high fall out rate, especially after first 6 months of treatment.

Several studies have shown that the individual HCV proteins are immunogenic in normal mice, including following immunisation with DNA. Several HCV vaccines are currently in clinical trial for either prophylaxis or therapy. The most advanced are currently in Phase 2 by Chiron and Innogenetics using E1 or E2 envelope proteins. An epitope vaccine by Transvax is also in Phase 2. Several vaccines are in preclinical development which use sequences from core and non-structural antigens using a variety of delivery systems including DNA.

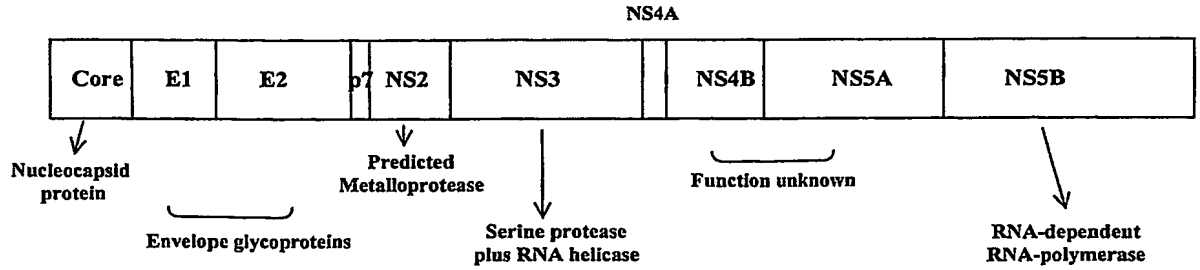
HCV is a positive strand RNA virus of the flaviviradae family, whose genome is 9.4kb in length, with one open reading frame. The HCV genome is translated as a single polyprotein, which is then processed by host and viral proteases to produce structural proteins (core, envelope E1 and E2) and six non-structural proteins with various enzymatic activities. The genome of the HCV J4L6 isolate, which is an example of the 1b genotype, is found as accession number AF054247 (Yanagi,M., St Claire,M., Shapiro,M., Emerson,S.U., Purcell,R.H. and Bukh,J. "Transcripts of a chimeric cDNA clone of hepatitis C virus genotype 1b are infectious in vivo". Virology 244 (1), 161-172 (1998)), and is shown in Figure 1.

The envelope proteins are responsible for recognition, binding and entry of virus onto target cells. The major non-structural proteins involved in viral replication include NS2 (Zn dependent metaloproteinase), NS3 (serine protease / helicase), NS4A (protease co-factor), NS5A and NS5B (RNA polymerase)(Bartenschlager B and Lohmann V. 2000. Replication of hepatitis C virus. J. Gen Virol 81, 1631).

The structure of the HCV polyprotein can be represented as follows (the figures refer to the position of the first amino acid of each protein; the full polyprotein of the J4L6 isolate is 3010 amino acids in length)

Co-ordinates of viral protein antigens:

Core	1-191
NS3	1027-1657
NS4B	1712-1972
NS5B	2420-3010



The virus has a high mutation rate and at least six major genotypes have been defined based in the nucleotide sequence of conserved and non-conserved regions. However there is additional heterogeneity as HCV isolated from a single patient is always presented as a mixture of closely related genomes or quasi-species.

The HCV genome shows a high degree of genetic variation, which has been classified into 6 major genotypes (1a, 1b, 2, 3, 4, 5,and 6). Genotypes 1a, 1b, 2 and 3 are the most prevalent in Europe, North and South America, Asia, China, Japan and Australia. Genotypes 4 and 5 are predominant in Africa and genotype 6 S.E Asia.

There is a great need, therefore, for improved treatments of HCV infection and also to provide treatments that are diverse in the ability to treat a number of HCV genotypes. In a first aspect of the present invention there is provided novel vaccine formulations that are diverse in their protection against various genotypes.

The vaccines of the present invention comprise oligonucleotides that encode the polypeptide sequences of one or more of the following HCV proteins: core, NS3, NS4B and NS5B. Preferably the vaccine comprises at least two HCV proteins selected from core, NS3, NS4 and NS5B; more preferably the vaccine comprises at least 3 HCV proteins selected from core, NS3, NS4 and NS5B; and most preferably the vaccine comprises core, NS3, NS4 and NS5B.

The oligonucleotide sequences used in the vaccines of the present invention are preferably DNA sequences.

The polypeptides encoded by the oligonucleotide vaccines of the present invention may comprise the full length amino acid sequence or alternatively the polypeptides may be shorter than the full length proteins, in that they comprise a sufficient proportion of the full length polynucleotide sequence to enable the expression product of the shortened gene to generate an immune response which cross reacts with the full length protein.

The oligonucleotide sequences encoding one or more of the selected HCV proteins are preferably codon optimised for expression in a mammalian cell.

In preferred vaccines of the present invention the HCV polypeptides are inactivated. For example the helicase and protease activity of NS3 is preferably reduced or abolished by mutation of the gene. Preferably NS5B polymerase activity of the expressed polypeptide is reduced or abolished by mutation. Preferably NS4B activity of the expressed polypeptide is reduced or abolished by mutation. Preferably activity of the Core protein of the expressed

polypeptide is reduced or abolished by mutation. Mutation in this sense could comprise an addition, deletion, substitution or rearrangement event to polynucleotide encoding the polypeptide. Alternatively the full length sequence may be expressed in two or more separate parts.

The functional structure and enzymatic function of the HCV polypeptides NS3 and NS5B are described in the art.

NS5B has been described as an RNA-dependent RNA polymerase Qin *et al.*, 2001, Hepatology, 33, pp 728-737; Lohmann *et al.*, 2000, Journal of Viral Hepatitis; Lohmann *et al.*, 1997, Nov., Journal of Virology, 8416-8428; De Francesco *et al.*, 2000, Seminars in Liver Disease, 20(1), 69-83. The NS5B polypeptide has been described as having four functional motifs A, B, C and D.

Preferably the NS5B polypeptide sequence encoded by polynucleotide vaccines of the present invention is mutated to reduce or remove RNA-dependent RNA polymerase activity. Preferably the polypeptide is mutated to disrupt motif A of NS5B, for example a substitution of the Aspartic acid (D) in position **2639** to Glycine (G); or a substitution of Aspartic acid (D) **2644** to Glycine (G). Preferably, the NS5B polypeptide encoded by the vaccine polynucleotide contains both of these Aspartic acid mutations.

Preferably, the encoded NS5B contains a disruption in its motif C. For example, Mutation of D_{2737}, an invariant aspartic acid residue, to H, N or E leads to the complete inactivation of NS5B.

Preferably the NS5B encoded by the DNA vaccines of the present invention comprise a motif A mutation, which may optionally comprise a motif C mutation.

NS3 has been described as having both protease and helicase activity. The NS3 polypeptides encoded by the DNA vaccines of the present invention are preferably mutated to disrupt both the protease and helicase activities of NS3. It is known that the protease activity of NS3 is linked to the "catalytic triad" of H-1083, D-1107 and S-1165. Preferably the NS3 encoded by the vaccines of the present invention comprises a mutation in the Catalytic triad residues, and most preferably the NS3 comprises single point mutation of Serine 1165 to valine (De Francesco, R., Pessi, a and Steinkuhler C. 1998. The hepatitis C Virus NS3 proteinase : structure and function of a zinc containing proteinase. Anti- Viral Therapy 3, 1-18.).

The structure and function of NS3 can be represented as:

Protease	Helicase

Catalytic triad:
H-1083
D-1107
S-1165

Established functional motifs:

I	II	III	IV
GKS	DECH	TAT	QRrGRtGR

Four critical motifs for the helicase activity of NS3 have been identified, I, II, III and IV. Preferably the NS3 encoded by the DNA vaccines of the present invention comprise disruptive mutations to at least one of these motifs. Most preferably, there is a substitution of the Aspartic acid 1316 to glutamine (Paolini, C, Lahm A, De Francesco R and Gallinari P 2000, Mutational analysis of hepatitis C virus NS3-associated helicase. J.Gen Virol. 81, 1649). Neither of these most preferred NS3 mutations, S1165V or D1316Q, lie within known or predicted T cell epitopes.

Most preferably the NS3 polypeptide encoded by the DNA vaccines of the present invention comprise Serine (S) 1165 to Valine (V) and an Aspartic acid (D) 1316 to Glutamine (Q) mutation.

The biological functions of HCV core protein are complex and do not correlate with discrete point mutations (McLauchlan J. 2000. Properties of the hepatitis C virus core protein: a structural protein that modulates cellular processes. J of Viral Hepatitis 7, 2-4). There is evidence that core directly interacts with the lymphotoxin β receptor, and can also interfere with NFκB and PKR pathways and can influence cell survival and apoptosis. A recombinant vaccinia construct expressing core was found to inhibit cellular responses to vaccinia making it more virulent in vivo.

The Core polypeptides used in the vaccines of the present invention are either full length or in a truncated form. Most preferably the Core polypeptide is full length, but the sequence of which is rearranged to abrogate any activity of Core protein. Preferably the Core polypeptide is split into at least two fragments, and most preferably forming a polypeptide consisting of Core amino acids 66-191 followed onto amino acids 1-65, and alternatively Core amino acids 105-191 followed by Core amino acids 1-104.

The sequence of full length NS4B is given in Figure X. The preferred polypeptide encoded by the polynucleotides of the present invention contain an N-terminal truncation to remove a region that is hypervariable between HCV isolates and genotypes. Preferably the

NS4B polypeptide contains a deletion of between 30– 100 amino acids from the N-terminus, more preferably between 40–80 amino acids, and most preferably a deletion of the first N-terminal 48 amino acids **(in the context of the J4 L6 isolate this corresponds to a truncation at** amino acid 1760, which is a loss of the first 48 amino acids of NS4B; equivalent truncations in other HCV isolates also form part of the present invention). Additionally, the NS4B sequence may be divided into two or more fragments and expressed in a polypeptide having the sequence of NS4B arranged in a different order to that found in the wild-type molecule.

The polynucleotides which are present in the vaccines of the present invention may comprise the natural nucleotide sequence as found in the HCV virus, however, it is preferred that the nucleotide sequence is codon optimised for expression in mammalian cells. Accordingly, in a second aspect of the present invention there is provided novel polynucleotides encoding HCV Core, NS3, NS4B and NS5B in which the codon usage closely reflects that of mammalian genes.

The codon usage in the polynucleotides of the present invention encoding HCV Core, NS3, NS4B and NS5B is preferably also altered such that rare codons do not appear in concentrated clusters, and are on the contrary relatively evenly spaced throughout the polynucleotide sequence.

The DNA code has 4 letters (A, T, C and G) and uses these to spell three letter "codons" which represent the amino acids of the proteins encoded in an organism's genes. The linear sequence of codons along the DNA molecule is translated into the linear sequence of amino acids in the protein(s) encoded by those genes. The code is highly degenerate, with 61 codons coding for the 20 natural amino acids and 3 codons representing "stop" signals. Thus, most amino acids are coded for by more than one codon - in fact several are coded for by four or more different codons.

Where more than one codon is available to code for a given amino acid, it has been observed that the codon usage patterns of organisms are highly non-random. Different species show a different bias in their codon selection and, furthermore, utilisation of codons may be markedly different in a single species between genes which are expressed at high and low levels. This bias is different in viruses, plants, bacteria and mammalian cells, and some species show a stronger bias away from a random codon selection than others. For example, humans and other mammals are less strongly biased than certain bacteria or viruses. For these

reasons, there is a significant probability that a mammalian gene expressed in E.coli or a viral gene expressed in mammalian cells will have an inappropriate distribution of codons for efficient expression. However, a gene with a codon usage pattern suitable for E.coli expression may also be efficiently expressed in humans. It is believed that the presence in a heterologous DNA sequence of clusters of codons which are rarely observed in the host in which expression is to occur, is predictive of low heterologous expression levels in that host.

There are several examples where changing codons from those which are rare in the host to those which are host-preferred ("codon optimisation") has enhanced heterologous expression levels, for example the BPV (bovine papilloma virus) late genes L1 and L2 have been codon optimised for mammalian codon usage patterns and this has been shown to give increased expression levels over the wild-type HPV sequences in mammalian (Cos-1) cell culture (Zhou et. al. J. Virol 1999. 73, 4972-4982). In this work, every BPV codon which occurred more than twice as frequently in BPV than in mammals (ratio of usage >2), and most codons with a usage ratio of >1.5 were conservatively replaced by the preferentially used mammalian codon. In WO97/31115, WO97/48370 and WO98/34640 (Merck & Co., Inc.) codon optimisation of HIV genes or segments thereof has been shown to result in increased protein expression and improved immunogenicity when the codon optimised sequences are used as DNA vaccines in the host mammal for which the optimisation was tailored. In these documents, the sequences consist entirely of optimised codons (except where this would introduce an undesired restriction site, intron splice site etc.) because each viral codon is conservatively replaced with the optimal codon for the intended host.

The term "codon usage pattern" refers to the average frequencies for all codons in the nucleotide sequence, gene or class of genes under discussion (e.g. highly expressed mammalian genes). Codon usage patterns for mammals, including humans can be found in the literature (see e.g. Nakamura et.al. Nucleic Acids Research 1996, 24:214-215).

In the polynucleotides of the present invention, the codon usage pattern is preferably altered from that typical of HCV to more closely represent the codon bias of the target organism, e.g. E.coli or a mammal, especially a human. The "codon usage coefficient" is a measure of how closely the codon usage pattern of a given polynucleotide sequence resembles that of a target species. The codon frequencies for each of the 61 codons (expressed as the number of occurrences per 1000 codons of the selected class of genes) are normalised for each of the twenty natural amino acids, so that the value for the most

frequently used codon for each amino acid is set to 1 and the frequencies for the less common codons are scaled to lie between zero and 1. Thus each of the 61 codons is assigned a value of 1 or lower for the highly expressed genes of the target species. In order to calculate a codon usage coefficient for a specific polynucleotide, relative to the highly expressed genes of that species, the scaled value for each codon of the specific polynucleotide are noted and the geometric mean of all these values is taken (by dividing the sum of the natural logs of these values by the total number of codons and take the anti-log). The coefficient will have a value between zero and 1 and the higher the coefficient the more codons in the polynucleotide are frequently used codons. If a polynucleotide sequence has a codon usage coefficient of 1, all of the codons are "most frequent" codons for highly expressed genes of the target species.

The present invention provides polynucleotide sequences which encode HCV Core, NS3, NS4B or NS5B amino acid sequences, wherein the codon usage pattern of the polynucleotide sequence resembles that of highly expressed mammalian genes. Preferably the polynucleotide sequence is a DNA sequence. Desirably the codon usage pattern of the polynucleotide sequence resembles that of highly expressed human genes.

The codon optimised polynucleotide sequence encoding HCV core (1-191) is shown in Figure 2. The codon optimised polynucleotide sequence encoding HCV NS3, comprising the S1165V and D1316Q polypeptide mutation, is shown in Figure 3. The codon optimised polynucleotide sequence encoding HCV NS4B, comprising the N terminal 1-48 truncation of the polypeptide, is shown in Figure 4. The codon optimised polynucleotide sequence encoding HCV NS5B, comprising the D2639G and D2644G polypeptide mutation, is shown in Figure 5.

Accordingly, there is provided a synthetic gene comprising a plurality of codons together encoding HCV Core, NS3, NS4B or NS5B amino acid sequences, wherein the selection of the possible codons used for encoding the amino acid sequence has been changed to resemble the optimal mammalian codon usage such that the frequency of codon usage in the synthetic gene more closely resembles that of highly expressed mammalian genes than that of Hepatitis C virus genes. Preferably the codon usage pattern is substantially the same as that for highly expressed human genes.

According to the present invention, the codon usage pattern of the polynucleotides will preferably exclude codons with an RSCU value of less than 0.2 in highly expressed genes of the target organism. A relative synonymous codon usage (RSCU) value is the

observed number of codons divided by the number expected if all codons for that amino acid were used equally frequently. A polynucleotide of the present invention will generally have a codon usage coefficient (as defined above) for highly expressed human genes of greater than 0.5, preferably greater than 0.6, most preferably greater than 0.7 but less than 1. Desirably the polynucleotide will also have a codon usage coefficient for highly expressed E.coli genes of greater than 0.5, preferably greater than 0.6, most preferably greater than 0.7.

The "natural" HCV core, NS3, NS4B and NS5B sequences have been analysed for codon usage. The Codon usage coefficient for the HCV proteins are Core (0.487), NS3 (0.482), NS4B-0.481 and NS5B (0.459).

The HCV core, NS3, NS4B and NS5B polynucleotides are also optimised to prevent clustering of rare, non-optimal, codons being present in concentrated areas. The polynucleotides, therefore, are optimised such that individual rare codons, such as those with a codon usage coefficient of <0.4 (and more preferably of <0.3) are evenly spaced throughout the polynucleotides.

Expression levels of codon optimised mutated Core, NS3 and NS5B have been shown to be increased compared to wild type, as assessed by Western blot. The truncated codon optimised NS4B has been expressed as a fusion with NS5B, and the fusion expresses well.

The vaccines of the present invention may comprise a vector that directs individual expression of the HCV polypeptides, alternatively all HCV polypeptides may be expressed as a large fusion protein.

The most preferred DNA vaccines of the present invention comprise fusion proteins. For example preferably the fusion protein comprises the nucleotide sequences for Core (1-191 (all but divide sequence into two or more fragments to disable biological activity)), NS3 1027-1657 (mutations to inactivate helicase (Aspartic acid 1316 to Glutamine) and protease (serine 1165 to valine) activity, NS5B 2420-3010 (mutation at Aspartic acid 2639 to Glycine and Aspartic acid 2644 to Glycine, Motif A) to inactivate polymerase activity), and NS4B 1712-1972 (optionally truncated to 1760-1972 remove N-terminal highly variable fragment). This fusion protein comprises approximately 1626 amino acids (4.9 kB).

Preferred fusion polyproteins include:

Polyprotein 1:

Core	NS3	NS4B	NS5B

Polyprotein 2:

NS3	NS4B	NS5B	Core

Polyprotein 3:

NS4B	NS5B	Core	NS3

Polyprotein 4:

NS5B	Core	NS3	NS4B

Polyprotein 5:

Core (66-191)-(1-65)	NS3	NS4B	NS5B

Polyprotein 6:

Core (105-191)-(1-104)	NS3	NS4B	NS5B

The present invention provides the novel DNA vaccines and polypeptides as described above. Also provided by the present invention are analogues of the described polypeptides and DNA vaccines comprising them.

The term "analogue" refers to a polynucleotide which encodes the same amino acid sequence as another polynucleotide of the present invention but which, through the redundancy of the genetic code, has a different nucleotide sequence whilst maintaining the same codon usage pattern, for example having the same codon usage coefficient or a codon usage coefficient within 0.1, preferably within 0.05 of that of the other polynucleotide.

Shorter polynucleotide sequences are within the scope of the invention. For example, a polynucleotide of the invention may encode a fragment of a HCV protein. A polynucleotide which encodes a fragment of at least 8, for example 8-10 amino acids or up to 20, 50, 60, 70, 80, 100, 150 or 200 amino acids in length is considered to fall within the scope of the invention as long as the polynucleotide has a codon usage pattern which resembles that of a highly expressed mammalian gene and the encoded oligo or polypeptide demonstrates HCV antigenicity. In particular, but not exclusively, this aspect of the invention encompasses the

situation when the polynucleotide encodes a fragment of a complete HCV protein sequence and may represent one or more discrete epitopes of that protein.

The polynucleotides according to the invention have utility in the production by expression of the encoded proteins, which expression may take place *in vitro, in vivo* or *ex vivo*. The nucleotides may therefore be involved in recombinant protein synthesis, for example to increase yields, or indeed may find use as therapeutic agents in their own right, utilised in DNA vaccination techniques. Where the polynucleotides of the present invention are used in the production of the encoded proteins *in vitro* or *ex vivo*, cells, for example in cell culture, will be modified to include the polynucleotide to be expressed. Such cells include transient, or preferably stable mammalian cell lines. Particular examples of cells which may be modified by insertion of vectors encoding for a polyproteins according to the invention include mammalian HEK293T, CHO, HeLa, 293 and COS cells. Preferably the cell line selected will be one which is not only stable, but also allows for mature glycosylation and cell surface expression of a polyprotein. Expression may be achieved in transformed oocytes. A polypeptide may be expressed from a polynucleotide of the present invention, in cells of a transgenic non-human animal, preferably a mouse. A transgenic non-human animal expressing a polypeptide from a polynucleotide of the invention is included within the scope of the invention.

The present invention includes expression vectors that comprise the nucleotide sequences of the invention. Such expression vectors are routinely constructed in the art of molecular biology and may for example involve the use of plasmid DNA and appropriate initiators, promoters, enhancers and other elements, such as for example polyadenylation signals which may be necessary, and which are positioned in the correct orientation, in order to allow for protein expression. Other suitable vectors would be apparent to persons skilled in the art. By way of further example in this regard we refer to Sambrook *et al.* Molecular Cloning: a Laboratory Manual. 2nd Edition. CSH Laboratory Press. (1989).

Preferably, a polynucleotide of the invention, or for use in the invention in a vector, is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence, such as a

promoter, "operably linked" to a coding sequence is positioned in such a way that expression of the coding sequence is achieved under conditions compatible with the regulatory sequence.

The vectors may be, for example, plasmids, artificial chromosomes (e.g. BAC, PAC, YAC), virus or phage vectors provided with a origin of replication, optionally a promoter for the expression of the polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin or kanamycin resistance gene in the case of a bacterial plasmid or a resistance gene for a fungal vector. Vectors may be used *in vitro*, for example for the production of DNA or RNA or used to transfect or transform a host cell, for example, a mammalian host cell e.g. for the production of protein encoded by the vector. The vectors may also be adapted to be used *in vivo*, for example in a method of DNA vaccination or of gene therapy.

Promoters and other expression regulation signals may be selected to be compatible with the host cell for which expression is designed. For example, mammalian promoters include the metallothionein promoter, which can be induced in response to heavy metals such as cadmium, and the β-actin promoter. Viral promoters such as the SV40 large T antigen promoter, human cytomegalovirus (CMV) immediate early (IE) promoter, rous sarcoma virus LTR promoter, adenovirus promoter, or a HPV promoter, particularly the HPV upstream regulatory region (URR) may also be used. All these promoters are well described and readily available in the art.

Examples of suitable viral vectors include herpes simplex viral vectors, vaccinia or alpha-virus vectors and retroviruses, including lentiviruses, adenoviruses and adeno-associated viruses. Gene transfer techniques using these viruses are known to those skilled in the art. Retrovirus vectors for example may be used to stably integrate the polynucleotide of the invention into the host genome, although such recombination is not preferred. Replication-defective adenovirus vectors by contrast remain episomal and therefore allow transient expression. Vectors capable of driving expression in insect cells (for example baculovirus vectors), in human cells or in bacteria may be employed in order to produce quantities of the HCV protein encoded by the polynucleotides of the present invention, for example for use as subunit vaccines or in immunoassays.

In a further aspect, the present invention provides a pharmaceutical composition comprising a polynucleotide sequence as described herein. Preferably the composition comprises a DNA vector according to the second aspect of the present invention. In preferred

embodiments the composition comprises a plurality of particles, preferably gold particles, coated with DNA comprising a vector encoding a polynucleotide sequence which encodes an HPV amino acid sequence, wherein the codon usage pattern of the polynucleotide sequence resembles that of highly expressed mammalian genes, particularly human genes. In alternative embodiments, the composition comprises a pharmaceutically acceptable excipient and a DNA vector according to the second aspect of the present invention. The composition may also include an adjuvant.

DNA vaccines may be delivered by interstitial administration of liquid vaccines into the muscle (WO90/11092) or by mechanisms other than intra-muscular injection. For example, delivery into the skin takes advantage of the fact that immune mechanisms are highly active in tissues that are barriers to infection such as skin and mucous membranes. Delivery into skin could be via injection, via jet injector (which forces a liquid into the skin, or underlying tissues including muscles, under pressure) or via particle bombardment, in which the DNA may be coated onto particles of sufficient density to penetrate the epithelium (US Patent No. 5371015). For example, the nucleotide sequences may be incorporated into a plasmid which is coated on to gold beads which are then administered under high pressure into the epidermis, such as, for example, as described in Haynes *et al* J. Biotechnology 44: 37-42 (1996). Projection of these particles into the skin results in direct transfection of both epidermal cells and epidermal Langerhan cells. Langerhan cells are antigen presenting cells (APC) which take up the DNA, express the encoded peptides, and process these for display on cell surface MHC proteins. Transfected Langerhan cells migrate to the lymph nodes where they present the displayed antigen fragments to lymphocytes, evoking an immune response. Very small amounts of DNA (less than 1μg, often less than 0.5μg) are required to induce an immune response via particle mediated delivery into skin and this contrasts with the milligram quantities of DNA known to be required to generate immune responses subsequent to direct intramuscular injection.

Where the polynucleotides of the present invention find use as therapeutic agents, e.g. in DNA vaccination, the nucleic acid will be administered to the mammal e.g. human to be vaccinated. The nucleic acid, such as RNA or DNA, preferably DNA, is provided in the form of a vector, such as those described above, which may be expressed in the cells of the mammal. The polynucleotides may be administered by any available technique. For example, the nucleic acid may be introduced by needle injection, preferably intradermally,

subcutaneously or intramuscularly. Alternatively, the nucleic acid may be delivered directly into the skin using a nucleic acid delivery device such as particle-mediated DNA delivery (PMDD). In this method, inert particles (such as gold beads) are coated with a nucleic acid, and are accelerated at speeds sufficient to enable them to penetrate a surface of a recipient (e.g. skin), for example by means of discharge under high pressure from a projecting device. (Particles coated with a nucleic acid molecule of the present invention are within the scope of the present invention, as are delivery devices loaded with such particles). The composition desirably comprises gold particles having an average diameter of 0.5-5μm, preferably about 2 μm. In preferred embodiments, the coated gold beads are loaded into tubing to serve as cartridges such that each cartridge contains 0.1-1 mg, preferably 0.5mg gold coated with 0.1-5 μg, preferably about 0.5 μg DNA/cartridge.

According to another aspect of the invention there is provided a host cell comprising a polynucleotide sequence as described herein. The host cell may be bacterial, e.g. E.coli, mammalian, e.g. human, or may be an insect cell. Mammalian cells comprising a vector according to the present invention may be cultured cells transfected in vitro or may be transfected in vivo by administration of the vector to the mammal.

In a further aspect, the present invention provides a method of making a pharmaceutical composition including the step of altering the codon usage pattern of a wild-type HCV nucleotide sequence, or creating a polynucleotide sequence synthetically, to produce a sequence having a codon usage pattern resembling that of highly expressed mammalian genes and encoding a wild-type HCV amino acid sequence or a mutated HCV amino acid sequence comprising the wild-type sequence with amino acid changes sufficient to inactivate one or more of the natural functions of the polypeptide.

Also provided are the use of a polynucleotide or vaccine as described herein, in the treatment or prophylaxis of an HCV infection.

Suitable techniques for introducing the naked polynucleotide or vector into a patient include topical application with an appropriate vehicle. The nucleic acid may be administered topically to the skin, or to mucosal surfaces for example by intranasal, oral, intravaginal or intrarectal administration. The naked polynucleotide or vector may be present together with a pharmaceutically acceptable excipient, such as phosphate buffered saline (PBS). DNA uptake may be further facilitated by use of facilitating agents such as bupivacaine, either separately or included in the DNA formulation. Other methods of administering the nucleic acid directly

to a recipient include ultrasound, electrical stimulation, electroporation and microseeding which is described in US-5,697,901.

Uptake of nucleic acid constructs may be enhanced by several known transfection techniques, for example those including the use of transfection agents. Examples of these agents includes cationic agents, for example, calcium phosphate and DEAE-Dextran and lipofectants, for example, lipofectam and transfectam. The dosage of the nucleic acid to be administered can be altered. Typically the nucleic acid is administered in an amount in the range of 1pg to 1mg, preferably 1pg to 10μg nucleic acid for particle mediated gene delivery and 10μg to 1mg for other routes.

A nucleic acid sequence of the present invention may also be administered by means of specialised delivery vectors useful in gene therapy. Gene therapy approaches are discussed for example by Verme *et al*, Nature 1997, 389:239-242. Both viral and non-viral vector systems can be used. Viral based systems include retroviral, lentiviral, adenoviral, adeno-associated viral, herpes viral, Canarypox and vaccinia-viral based systems. Non-viral based systems include direct administration of nucleic acids, microsphere encapsulation technology (poly(lactide-co-glycolide) and, liposome-based systems. Viral and non-viral delivery systems may be combined where it is desirable to provide booster injections after an initial vaccination, for example an initial "prime" DNA vaccination using a non-viral vector such as a plasmid followed by one or more "boost" vaccinations using a viral vector or non-viral based system.

A nucleic acid sequence of the present invention may also be administered by means of transformed cells. Such cells include cells harvested from a subject. The naked polynucleotide or vector of the present invention can be introduced into such cells *in vitro* and the transformed cells can later be returned to the subject. The polynucleotide of the invention may integrate into nucleic acid already present in a cell by homologous recombination events. A transformed cell may, if desired, be grown up *in vitro* and one or more of the resultant cells may be used in the present invention. Cells can be provided at an appropriate site in a patient by known surgical or microsurgical techniques (e.g. grafting, micro-injection, etc.)

Suitable cells include antigen-presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for

presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumour, e.g. anti-cervical carcinoma effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumour and peri-tumoural tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells, either for transformation in vitro and return to the patient or as the in vivo target of nucleotides delivered in the vaccine, for example by particle mediated DNA delivery. Dendritic cells are highly potent APCs (*Banchereau* and Steinman, *Nature 392*:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumour immunity (*see* Timmerman and Levy, *Ann. Rev. Med. 50*:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, for example the antigen(s) encoded in the constructs of the invention, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel et al., *Nature Med. 4*:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumour-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF, CD40 ligand, lipopolysaccharide LPS, flt3 ligand (a cytokine important in the generation of professional antigen presenting cells, particularly dendritic cells) and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

APCs may generally be transfected with a polynucleotide encoding an antigenic HCV amino acid sequence, such as a codon-optimised polynucleotide as envisaged in the present

invention. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and ex *vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the particle mediated approach described by Mahvi et al., *Immunology and cell Biology 75*:456-460, 1997.

Vaccines and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a vaccine or pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use. Vaccines comprising nucleotide sequences intended for administration via particle mediated delivery may be presented as cartridges suitable for use with a compressed gas delivery instrument, in which case the cartridges may consist of hollow tubes the inner surface of which is coated with particles bearing the vaccine nucleotide sequence, optionally in the presence of other pharmaceutically acceptable ingredients.

The pharmaceutical compositions of the present invention may include adjuvant compounds, or other substances which may serve to modulate or increase the immune response induced by the protein which is encoded by the DNA. These may be encoded by the DNA, either separately from or as a fusion with the antigen, or may be included as non-DNA elements of the formulation. Examples of adjuvant-type substances which may be included in the formulations of the present invention include ubiquitin, lysosomal associated membrane protein (LAMP), hepatitis B virus core antigen, flt3-ligand and other cytokines such as IFN-γ and GMCSF.

Other suitable adjuvants are commercially available such as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Imiquimod (3M, St. Paul, MN); Resimiquimod (3M, St. Paul, MN); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); aluminium salts such as aluminium hydroxide gel (alum) or aluminium phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides;

polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

In the formulations of the invention it is preferred that the adjuvant composition induces an immune response predominantly of the Th1 type. Thus the adjuvant may serve to modulate the immune response generated in response to the DNA-encoded antigens from a predominantly Th2 to a predominantly Th1 type response. High levels of Th1-type cytokines (*e.g.*, IFN-, TNF, IL-2 and IL-12) tend to favour the induction of cell mediated immune responses to an administered antigen. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol. 7*:145-173, 1989.

Accordingly, suitable adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL) together with an aluminium salt. Other known adjuvants which preferentially induce a TH1 type immune response include CpG containing oligonucleotides. The oligonucleotides are characterised in that the CpG dinucleotide is unmethylated. Such oligonucleotides are well known and are described in, for example WO96/02555. Immunostimulatory DNA sequences are also described, for example, by Sato et al., *Science 273*:352, 1996. CpG-containing oligonucleotides may be encoded separately from the papilloma antigen(s) in the same or a different polynucleotide construct, or may be immediately adjacent thereto, e.g. as a fusion therewith. Alternatively the CpG-containing oligonucleotides may be administered separately i.e. not as part of the composition which includes the encoded antigen. CpG oligonucleotides may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG and QS21 as disclosed in WO 00/09159 and WO 00/62800. Preferably the formulation additionally comprises an oil in water emulsion and/or tocopherol.

Another preferred adjuvant is a saponin, preferably QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO

94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), Detox (Ribi, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs).

Where the vaccine includes an adjuvant, the vaccine formulation may be administered in two parts. For example, the part of the formulation containing the nucleotide construct which encodes the antigen may be administered first, e.g. by subcutaneous or intramuscular injection, or by intradermal particle-mediated delivery, then the part of the formulation containing the adjuvant may be administered subsequently, either immediately or after a suitable time period which will be apparent to the physician skilled in the vaccines arts. Under these circumstances the adjuvant may be administered by the same route as the antigenic formulation or by an alternate route. In other embodiments the adjuvant part of the formulation will be administered before the antigenic part. In one embodiment, the adjuvant is administered as a topical formulation, applied to the skin at the site of particle mediated delivery of the nucleotide sequences which encode the antigen(s), either before or after the particle mediated delivery thereof.

Example 1,

Mutations introduced into antigen panel :-

1). Consensus mutations

A comparison of the full genome sequences of all known HCV isolates was carried out. Certain positions within the J4L6 polyprotein were identified as unusual/ deviating from the majority of other HCV isolates. With particular importance were those positions found to deviate from a more consensus residue across related 1b-group isolates, extending across groups 1a, 2, 3, and others, where one or two alternative amino acid residues otherwise

dominated in the equivalent position. None of the chosen consensus mutations interferes with a known CD4 or CD8 epitope. Two changes within NS3 actually restore an immunodominant HLA-B35-restricted CD8 epitope [Isoleucine (I) 1365 to Valine (V) and Glycine (G) 1366 to Alanine (A)].

The first 51 amino acids of NS4B have been removed due to unuseful variability.

Core

Alanine (A) 52 to Threonine (T)

NS3

Valine (V) 1040 to Leucine (L)
Leucine (L) 1106 to Glutamine (Q)
Serine (S) 1124 to Threonine (T)
Valine (V) 1179 to Isoleucine (I)
Threonine (T) 1215 to Serine (S)
Glycine (G) 1289 to Alanine (A)
Serine (S) 1290 to Proline (P)
Isoleucine (I) 1365 to Valine (V)
Glycine (G) 1366 to Alanine (A)
Threonine (T) 1408 to Serine (S)
Proline (P) 1428 to Threonine (T)
Isoleucine (I) 1429 to Serine (S)
Isoleucine (I) 1636 to Threonine (T)

NS4B

Start ORF at Phenylalanine (F) 1760

NS5B

Isoleucine (I) 2824 to Valine (V)

Threonine (T) 2892 to Serine (S)

Threonine (T) 2918 to Valine (V)

N.B. Numbering is according to position in polyprotein for J4L6 isolate.

Example 2, *Construction of plasmid DNA vaccines*

Polynucleotide sequences encoding HCV Core, NS3, truncated NS4B, and NS5B, were codon optimised for mammalian codon usage using SynGene 2e software. The codon usage coefficient was improved to greater than 0.7 for each polynucleotide.

The sense and anti-sense strands of each new polynucleotide sequence, incorporating codon optimisation, enzymatic knockout mutations, and consensus mutations, were divided into regions of 40-60 nucleotides, with a 20 nucleotide overlap. These regions were synthesised commercially and the polynucleotide generated by an oligo assembly PCR method.

The outer forward and reverse PCR primers for each polynucleotide, illustrating unique restriction endonuclease sites used for cloning, are outlined below:

HCV Core

Forward primer

5'-GAATTC**GCGGCCGCCATG**AGCACCAACCCCAAGCCCCAGCGCAAGACCAAGCGGAACACC-3'

NotI translation start codon

Reverse primer

5'-GAATTC**GGATCCTCA**TGCGCTAGCGGGGATGGTGAGGCAGCTCAGCAGCGCCAGCAGGA-3'

BamHI Stop codon

HCV NS3

Forward primer

5'-GAATTC**GCGGCCGCCATG**GCCCCCATCACCGCCTACAGCCAGCAGACCCGGGGAC-3'

NotI translation start codon

R verse primer

5'-GAATTC**GGATCCTCA**GGTGACCACCTCCAGGTCAGCGGACATGCACGCCATGATG-3'

BamHI Stop codon

HCV NS4B
Forward primer
5'-GAATTC**GCGGCCGC**C**ATG**TTTTGGGCCAAGCATATGTGGAACTTCA-3'
NotI translation start codon

Reverse primer
5'-GAATTC**GGATCCTCA**GCAAGGGGTGGAGCAGTCCTCGTTGATCCAC-3'
BamHI Stop codon

HCV NS5B
Forward primer
5'-GAATTC**GCGGCCGC**C**ATG**TCCATGTCCTACACCTGGACCGGCGCCCTGA-3'
NotI translation start codon
Reverse primer
5'-GAATTC**GGATCCTCA**GCGGTTGGGCAGCAGGTAGATGCCGACTCCGACG-3'
BamHI Stop codon

All polynucleotides, encoding single antigens, were cloned into mammalian expression vector p7313ie via Not I and BamHI unique cloning sites (see figure 7).

The polyproteins that were encoded were as follows (including mutations and codon optimisations):

HCV Core translation:

MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERS
QPRGRRQPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRGSRPSWGPTDP
RRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGN
LPGCSFSIFLLALLSCLTIPASA

HCV NS3 translation:

MAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVSTATQSFLATCINGVCWTVY
HGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSMTPCTCGSSDLYLVTRHA
DVIPVRRRGDSRGSLLSPRPVSYLKGSVGGPLLCPSGHVVGIFRAAVCTRGVAKAVD
FIPVESMETTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAQGYKV
LVLNPSVAATLGFGAYMSKAHGIDPNIRTGVRTITTGAPITYSTYGKFLADGGCSGGA
YDIIICQECHSTDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSN
NGEIPFYGKAIPIEAIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAYYRGLDVSVIPT
SGDVVVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRS
QRRGRTGRGRSGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRLRAY
LNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQ
APPPSWDQMWKCLIRLKPTLHGPTPLLYRLGAVQNEVTLTHPITKYIMACMSADLEV
VT

HCV NS4B translation:

MFWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTQNTLLFNILGGWV
AAQLAPPSAASAFVGAGIAGAAVGSIGLGKVLVDILAGYGAGVAGALVAFKVMSGE
VPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGNH
VSPTHYVPESDAAARVTQILSSLTITQLLKRLHQWINEDCSTPC

HCV NS5B translation:

MSMSYTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDR
LQVLDDHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLS
SRAVNHIRSVWEDLLEDTETPIDTTIMAKSEVFCVQPEKGGRKPARLIVFPDLGVRVC
EKMALYDVVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYGTRCFG
STVTESDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASG
VLTTSCGNTLTCYLKATAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAAALRAF
TEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAA
WETARHTPVNSWLGNIIMYAPTLWARMILMTHFFSILLAQEQLEKALDCQIYGACYS
IEPLDLPQIIERLHGLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLL
SQGGRAATCGRYLFNWAVRTKLKLTPIPAASQLDLSGWFVAGYSGGDIYHSLSRAR
PRWFPLCLLLLSVGVGIYLLPNR

Example 3, *Immune response assays*

C57BL or BALB/c mice were immunised with either WT or codon optimised + mutated versions of the four HCV antigens expressed individually in the p7313 vector. Mice were immunised by PMID with a standard dose of 1.0 μg/cartridge and boosted and day 21 (boost 1), and again at day 49 (boost 2). Spleen cells were harvested from individual mice and restimulated in ELISPOT with different HCV antigen preparations. Both IL2 and IFNγ responses were measured. The reagents used to measure immune responses were purified HCV core, NS3, NS4 and NS5B (genotype 1b) proteins from Mikrogen, Vacinnia-Core and Vaccinia NS3-5 (genotype 1b in house).

HCV Core

C57BL Mice immunised with WT full length (FL-1-191) or truncated (TR 1-115) core were restimulated with HCV core protein and good responses were observed with purified core protein (figure 8)

HCV NS3

Mice were immunised with p7313 WT and codon optimised NS3 using PMID. Good responses to NS3 following immunisation and a single boost were demonstrated in C57Bl mice using both NS3 protein and Vaccinia 3-5 to read out the response by ELISPOT. Both IL2 and IFNγ responses were detected. No significant differences between wild type and codon optimised (co + m) versions of the constructs were observed in this experiment (figure 9). However differences in *in vitro* expression following transient transfection were observed between wild type and codon optimised constructs. Experiments to compare constructs at lower DNA dose or in the primary response may reveal differences in the potency of the plasmids.

HCV NS4B

Responses to full length WT p7313 NS4B were observed following PMID immunisation of BALB/c mice. Both IL2 and IFNγ ELISPOT responses were observed following in vitro restimulation with either NS4B protein and Vaccinia 3-5 (figure 10).

The NS4B protein was truncated at the N-terminus to remove a highly variable region, however expression of this protein could not be detected following *in vitro* tranfection studies because the available anti-sera had been raised against the N-terminal region. In order to confirm expression of this region it was fused with the NS5B protein. Recent experiments have confirmed that immune responses can be detected against the truncated NS4B protein, either alone or as a fusion with NS5B, using the NS4B protein and NS3-5 vaccinia. Good responses were observed to WT and codon optimised NS4B.

HCV NS5B

The immune response to NS5B following PMID was investigated following immunisation with WT and codon optimised (co + M) sequences. Good responses to NS5B following immunisation and a single boost were demonstrated in C57BL mice using both NS3 protein and vaccinia 3-5 to read out the response by ELISPOT. As with NS3 no differences in the immune response were observed between WT and co +m versions of the constructs in this experiment (figure 11).

Claims

1. An HCV vaccine of the present invention comprise oligonucleotides that encode the polypeptide sequences of one or more of the following HCV proteins: core, NS3, NS4B and NS5B.

2. An HCV vaccine as claimed in claim 1 wherein vaccine comprises at least two HCV proteins selected from core, NS3, NS4 and NS5B.

3. An HCV vaccine as claimed in claim 1 wherein the vaccine comprises at least 3 HCV proteins selected from core, NS3, NS4 and NS5B

4. An HCV vaccine as claimed in claim 1 wherein the vaccine comprises core, NS3, NS4 and NS5B.

5. A vaccine as claimed in any one of claims 1 to 4 wherein the oligonucleotides are codon optimised for expression in mammalian cells

6. A method of preventing or treating an HCV infection in a mammal comprising administering a vaccine as claimed in any one of claims 1 to 5 to a mammal.

Figure 1, HCV J4L6 genome wild-type cDNA sequence, reference accession number AF054247,

```
   1 gccagccccc tgatgggggc gacactccac catgaatcac tcccctgtga ggaactactg
  61 tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgtcgtgcag cctccaggac
 121 ccccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag
 181 gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttgg gcgtgccccc
 241 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg
 301 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac
 361 ctcaaagaaa aaccaaacgt aacaccaacc gccgcccaca ggacgtcaag ttcccgggcg
 421 gtggtcagat cgttggtgga gtttacctgt tgccgcgcag gggccccagg ttgggtgtgc
 481 gcgcgactag gaaggcttcc gagcggtcgc aacctcgtgg aaggcgacaa cctatcccaa
 541 aggctcgccg acccgagggc agggcctggg ctcagcccgg gtaccttgg ccctctatg
 601 gcaatgaggg cctggggtgg gcaggatggc tcctgtcacc ccgcggctcc cggcctagtt
 661 ggggccccac ggacccccgg cgtaggtcgc gtaacttggg taaggtcatc gataccctta
 721 catgcggctt cgccgatctc atggggtaca ttccgctcgt cggcgccccc ctaggggcg
 781 ctgccagggc cttggcacac ggtgtccggg ttctggagga cggcgtgaac tatgcaacag
 841 ggaacttgcc cggttgctct ttctctatct tcctcttggc tctgctgtcc tgtttgacca
 901 tcccagcttc cgcttatgaa gtgcgcaacg tgtccgggat ataccatgtc acgaacgact
 961 gctccaactc aagcattgtg tatgaggcag cggacgtgat catgcatact cccgggtgcg
1021 tgccctgtgt tcaggagggt aacagctccc gttgctgggt agcgctcact cccacgctcg
1081 cggccaggaa tgccagcgtc cccactacga caatacgacg ccacgtcgac ttgctcgttg
1141 ggacggctgc tttctgctcc gctatgtacg tgggggatct ctgcggatct attttcctcg
1201 tctcccagct gttcaccttc tcgcctcgcc ggcatgagac agtgcaggac tgcaactgct
1261 caatctatcc cggccatgta tcaggtcacc gcatggcttg ggatatgatg atgaactggt
1321 cacctacaac agccctagtg gtgtcgcagt tgctccggat cccacaagct gtcgtggaca
1381 tggtggcggg ggcccactgg ggagtcctgg cgggccttgc ctactattcc atggtaggga
1441 actgggctaa ggttctgatt gtggcgctac tctttgccgg cgttgacggg gagacccaca
1501 cgacggggag ggtggccggc cacaccacct ccgggttcac gtccctttc tcatctgggg
1561 cgtctcagaa aatccagctt gtgaatacca acggcagctg gcacatcaac aggactgccc
1621 taaattgcaa tgactccctc caaactgggt tctttgccgc gctgttttac gcacacaagt
1681 tcaactcgtc cgggtgcccg gagcgcatgg ccagctgccg ccccattgac tggttcgccc
1741 aggggtgggg ccccatcacc tatactaagc ctaacagctc ggatcagagg ccttattgct
1801 ggcattacgc gcctcgaccg tgtggtgtcg tacccgcgtc gcaggtgtgt ggtccagtgt
1861 attgtttcac cccaagccct gttgtggtgg ggaccaccga tcgttccggt gtccctacgt
1921 atagctgggg ggagaatgag acagacgtga tgctcctcaa caacacgcgt ccgccacaag
1981 gcaactggtt cggctgtaca tggatgaata gtactgggtt cactaagacg tgcggaggtc
2041 ccccgtgtaa catcgggggg gtcggtaacc gcaccttgat ctgccccacg gactgcttcc
2101 ggaagcaccc cgaggctact tacacaaaat gtggctcggg gccctggttg acacctaggt
```

2161 gcctagtaga ctacccatac aggctttggc actacccctg cactctcaat ttttccatct
2221 ttaaggttag gatgtatgtg gggggcgtgg agcacaggct caatgccgca tgcaattgga
2281 ctcgaggaga gcgctgtaac ttggaggaca gggataggtc agaactcagc ccgctgctgc
2341 tgtctacaac agagtggcag atactgccct gtgctttcac caccctaccg gctttatcca
2401 ctggtttgat ccatctccat cagaacatcg tggacgtgca atacctgtac ggtgtagggt
2461 cagcgtttgt ctcctttgca atcaaatggg agtacatcct gttgcttttc cttctcctgg
2521 cagacgcgcg cgtgtgtgcc tgcttgtgga tgatgctgct gatagcccag gctgaggccg
2581 ccttagagaa cttggtggtc ctcaatgcgg cgtccgtggc cggagcgcat ggtattctct
2641 cctttcttgt gttcttctgc gccgcctggt acattaaggg caggctggct cctggggcgg
2701 cgtatgcttt ttatggcgta tggccgctgc tcctgctcct actggcgtta ccaccacgag
2761 cttacgcctt ggaccgggag atggctgcat cgtgcggggg tgcggttctt gtaggtctgg
2821 tattcttgac cttgtcacca tactacaaag tgtttctcac taggctcata tggtggttac
2881 aatactttat caccagagcc gaggcgcaca tgcaagtgtg ggtccccccc ctcaacgttc
2941 ggggaggccg cgatgccatc atcctcctca cgtgtgcggt tcatccagag ttaatttttg
3001 acatcaccaa actcctgctc gccatactcg gcccgctcat ggtgctccag gctggcataa
3061 cgagagtgcc gtacttcgtg cgcgctcaag ggctcattcg tgcatgcatg ttagtgcgaa
3121 aagtcgccgg gggtcattat gtccaaatgg tcttcatgaa gctgggcgcg ctgacaggta
3181 cgtacgttta taaccatctt accccactgc gggactgggc ccacgcgggc ctacgagacc
3241 ttgcggtggc ggtagagccc gtcgtcttct ccgccatgga gaccaaggtc atcacctggg
3301 gagcagacac cgctgcgtgt ggggacatca tcttgggtct acccgtctcc gcccgaaggg
3361 ggaaggagat atttttggga ccggctgata gtctcgaagg gcaagggtgg cgactccttg
3421 cgcccatcac ggcctactcc caacaaacgc ggggcgtact tggttgcatc atcactagcc
3481 tcacaggccg ggacaagaac caggtcgaag gggaggttca agtggtttct accgcaacac
3541 aatctttcct ggcgacctgc atcaacggcg tgtgctggac tgtctaccat ggcgctggct
3601 cgaagaccct agccggtcca aaaggtccaa tcacccaaat gtacaccaat gtagacctgg
3661 acctcgtcgg ctggcaggcg ccccccgggg cgcgctccat gacaccatgc agctgtggca
3721 gctcggacct ttacttggtc acgagacatg ctgatgtcat tccggtgcgc cggcgaggcg
3781 acagcagggg aagtctactc tcccccaggc ccgtctccta cctgaaaggc tcctcgggtg
3841 gtccattgct ttgcccttcg gggcacgtcg tgggcgtctt ccgggctgct gtgtgcaccc
3901 ggggggtcgc gaaggcggtg gacttcatac ccgttgagtc tatggaaact accatgcggt
3961 ctccggtctt cacagacaac tcaacccccc cggctgtacc gcagacattc caagtggcac
4021 atctgcacgc tcctactggc agcggcaaga gcaccaaagt gccggctgcg tatgcagccc
4081 aagggtacaa ggtgctcgtc ctgaacccgt ccgttgccgc caccttaggg tttggggcgt
4141 atatgtccaa ggcacacggt atcgacccta acatcagaac tggggtaagg accattacca
4201 cgggcggctc cattacgtac tccacctatg gcaagttcct tgccgacggt ggctgttctg
4261 ggggcgccta tgacatcata atatgtgatg agtgccactc aactgactcg actaccatct
4321 tgggcatcgg cacagtcctg gaccaagcgg agacggctgg agcgcggctc gtcgtgctcg
4381 ccaccgctac acctccggga tcggttaccg tgccacaccc caatatcgag gaaataggcc
4441 tgtccaacaa tggagagatc cccttctatg gcaaagccat ccccattgag gccatcaagg

4501 gggggaggca tctcattttc tgccattcca agaagaaatg tgacgagctc gccgcaaagc
4561 tgacaggcct cggactgaac gctgtagcat attaccgggg ccttgatgtg tccgtcatac
4621 cgcctatcgg agacgtcgtt gtcgtggcaa cagacgctct aatgacgggt ttcaccggcg
4681 attttgactc agtgatcgac tgcaatacat gtgtcaccca gacagtcgac ttcagcttgg
4741 atcccacctt caccattgag acgacgaccg tgccccaaga cgcggtgtcg cgctcgcaac
4801 ggcgaggtag aactggcagg ggtaggagtg gcatctacag gtttgtgact ccaggagaac
4861 ggccctcggg catgttcgat tcttcggtcc tgtgtgagtg ctatgacgcg ggctgtgctt
4921 ggtatgagct cacgcccgct gagacctcgg ttaggttgcg ggcttaccta aatacaccag
4981 ggttgcccgt ctgccaggac catctggagt tctgggagag cgtcttcaca ggcctcaccc
5041 acatagatgc ccacttcctg tcccagacta aacaggcagg agacaacttt ccttacctgg
5101 tggcatatca agctacagtg tgcgccaggg ctcaagctcc acctccatcg tgggaccaaa
5161 tgtggaagtg tctcatacgg ctgaaaccta cactgcacgg gccaacaccc ctgctgtata
5221 ggctaggagc cgtccaaaat gaggtcatcc tcacacaccc cataactaaa tacatcatgg
5281 catgcatgtc ggctgacctg gaggtcgtca ctagcacctg ggtgctggta ggcggagtcc
5341 ttgcagcttt ggccgcatac tgcctgacga caggcagtgt ggtcattgtg ggcaggatca
5401 tcttgtccgg gaagccagct gtcgttcccg acagggaagt cctctaccag gagttcgatg
5461 agatggaaga gtgtgcctca caacttcctt acatcgagca gggaatgcag ctcgccgagc
5521 aattcaagca aaaggcgctc gggttgttgc aaacggccac caagcaagcg gaggctgctg
5581 ctcccgtggt ggagtccaag tggcgagccc ttgagacctt ctgggcgaag cacatgtgga
5641 atttcatcag cggaatacag tacctagcag gcttatccac tctgcctgga aaccccgcga
5701 tagcatcatt gatggcattt acagcttcta tcactagccc gctcaccacc caaaacaccc
5761 tcctgtttaa catcttgggg ggatgggtgg ctgcccaact cgctcctccc agcgctgcgt
5821 cagctttcgt gggcgccggc atcgccggag cggctgttgg cagcataggc cttgggaagg
5881 tgctcgtgga catcttggcg ggctatgggg caggggtagc cggcgcactc gtggccttta
5941 aggtcatgag cggcgaggtg ccctccaccg aggacctggt caacttactc cctgccatcc
6001 tctctcctgg tgccctggtc gtcggggtcg tgtgcgcagc aatactgcgt cggcacgtgg
6061 gcccgggaga gggggctgtg cagtggatga accggctgat agcgttcgct tcgcggggta
6121 accacgtctc ccctacgcac tatgtgcctg agagcgacgc tgcagcacgt gtcactcaga
6181 tcctctctag ccttaccatc actcaactgc tgaagcggct ccaccagtgg attaatgagg
6241 actgctctac gccatgctcc ggctcgtggc taagggatgt ttgggattgg atatgcacgg
6301 tgttgactga cttcaagacc tggctccagt ccaaactcct gccgcggtta ccgggagtcc
6361 ctttcctgtc atgccaacgc gggtacaagg gagtctggcg gggggacggc atcatgcaaa
6421 ccacctgccc atgcggagca cagatcgccg gacatgtcaa aaacggttcc atgaggatcg
6481 tagggcctag aacctgcagc aacacgtggc acggaacgtt ccccatcaac gcatacacca
6541 cgggaccttg cacaccctcc ccggcgccca actattccag ggcgctatgg cgggtggctg
6601 ctgaggagta cgtggaggtt acgcgtgtgg gggatttcca ctacgtgacg ggcatgacca
6661 ctgacaacgt aaagtgccca tgccaggttc cggcccccga attcttcacg gaggtggatg
6721 gagtgcggtt gcacaggtac gctccggcgt gcaaacctct tctacgggag gacgtcacgt
6781 tccaggtcgg gctcaaccaa tacttggtcg ggtcgcagct cccatgcgag cccgaaccgg

6841 acgtaacagt gcttacttcc atgctcaccg atccctccca cattacagca gagacggcta
6901 agcgtaggct ggctagaggg tctcccccct ctttagccag ctcatcagct agccagttgt
6961 ctgcgccttc tttgaaggcg acatgcacta cccaccatga ctccccggac gctgacctca
7021 tcgaggccaa cctcttgtgg cggcaggaga tgggcggaaa catcactcgc gtggagtcag
7081 agaataaggt agtaattctg gactctttcg aaccgcttca cgcggagggg gatgagaggg
7141 agatatccgt cgcggcggag atcctgcgaa aatccaggaa gttcccctca gcgttgccca
7201 tatgggcacg cccggactac aatcctccac tgctagagtc ctggaaggac ccggactacg
7261 tccctccggt ggtacacgga tgcccattgc cacctaccaa ggctcctcca ataccacctc
7321 cacggagaaa gaggacggtt gtcctgacag aatccaatgt gtcttctgcc ttggcggagc
7381 tcgccactaa gaccttcggt agctccggat cgtcggccgt tgatagcggc acggcgaccg
7441 cccttcctga cctggcctcc gacgacggtg acaaaggatc cgacgttgag tcgtactcct
7501 ccatgccccc ccttgaaggg gagccggggg accccgatct cagcgacggg tcttggtcta
7561 ccgtgagtga ggaggctagt gaggatgtcg tctgctgctc aatgtcctat acgtggacag
7621 gcgccctgat cacgccatgc gctgcggagg aaagtaagct gcccatcaac ccgttgagca
7681 actctttgct gcgtcaccac aacatggtct acgccacaac atcccgcagc gcaagcctcc
7741 ggcagaagaa ggtcaccttt gacagattgc aagtcctgga tgatcattac cgggacgtac
7801 tcaaggagat gaaggcgaag gcgtccacag ttaaggctaa gcttctatct atagaggagg
7861 cctgcaagct gacgccccca cattcggcca aatccaaatt tggctatggg gcaaaggacg
7921 tccggaacct atccagcagg gccgttaacc acatccgctc cgtgtgggag gacttgctgg
7981 aagacactga aacaccaatt gacaccacca tcatggcaaa aagtgaggtt ttctgcgtcc
8041 aaccagagaa gggaggccgc aagccagctc gccttatcgt attcccagac ctgggagttc
8101 gtgtatgcga gaagatggcc ctttacgacg tggtctccac ccttcctcag gccgtgatgg
8161 gctcctcata cggatttcaa tactccccca agcagcgggt cgagttcctg gtgaatacct
8221 ggaaatcaaa gaaatgccct atgggcttct catatgacac ccgctgtttt gactcaacgg
8281 tcactgagag tgacattcgt gttgaggagt caatttacca atgttgtgac ttggcccccg
8341 aggccagaca ggccataagg tcgctcacag agcggcttta catcgggggt cccctgacta
8401 actcaaaagg gcagaactgc ggttatcgcc ggtgccgcgc aagtggcgtg ctgacgacta
8461 gctgcggtaa taccctcaca tgttacttga aggccactgc agcctgtcga gctgcaaagc
8521 tccaggactg cacgatgctc gtgaacggag acgaccttgt cgttatctgt gaaagcgcgg
8581 gaacccagga ggatgcggcg gccctacgag ccttcacgga ggctatgact aggtattccg
8641 ccccccccgg ggatccgccc caaccagaat acgacctgga gctgataaca tcatgttcct
8701 ccaatgtgtc agtcgcgcac gatgcatctg gcaaaagggt atactacctc acccgtgacc
8761 ccaccacccc ccttgcacgg gctgcgtggg agacagctag acacactcca atcaactctt
8821 ggctaggcaa tatcatcatg tatgcgccca ccctatgggc aaggatgatt ctgatgactc
8881 acttttctc catccttcta gctcaagagc aacttgaaaa agccctggat tgtcagatct
8941 acggggcttg ctactccatt gagccacttg acctacctca gatcattgaa cgactccatg
9001 gtcttagcgc atttacactc cacagttact ctccaggtga gatcaatagg gtggcttcat
9061 gcctcaggaa acttggggta ccacccttgc gaacctggag acatcgggcc agaagtgtcc
9121 gcgctaagct actgtcccag gggggggaggg ccgccacttg tggcagatac ctctttaact

9181 gggcagtaag gaccaagctt aaactcactc caatcccggc cgcgtcccag ctggacttgt
9241 ctggctggtt cgtcgctggt tacagcgggg gagacatata tcacagcctg tctcgtgccc
9301 gaccccgctg gtttccgttg tgcctactcc tactttctgt aggggtaggc atttacctgc
9361 tccccaaccg atgaacgggg agctaaccac tccaggcctt aagccatttc ctgttttttt
9421 ttttttttt ttttttttt tctttttttt tttctttcct ttccttcttt ttttcctttc
9481 tttttccctt ctttaatggt ggctccatct tagccctagt cacggctagc tgtgaaaggt
9541 ccgtgagccg catgactgca gagagtgctg atactggcct ctctgcagat catgt

Figure 2, codon optimised HCV Core polynucleotide

ATGAGCACCAACCCCAAGCCCCAGCGCAAGACCAAGCGGAACACCAACCGGA
GACCCCAGGACGTCAAGTTCCCAGGAGGAGGCCAGATCGTGGGCGGCGTGTA
CCTGCTGCCCCGCCGGGGGCCCCGGCTGGGCGTGCGCGCCACCCGCAAGAC
CAGCGAGCGCTCCCAGCCAAGAGGCAGACGCCAGCCGATCCCGAAGGCCCGC
CGCCCTGAGGGCCGGGCTTGGGCCCAGCCAGGCTACCCCTGGCCCCTGTATG
GCAACGAGGGCCTGGGATGGGCTGGGTGGCTCCTCAGCCCCCGGGGGTCTAG
GCCCAGTTGGGGACCGACCGACCCCCCGCAGGCGCAGCCGCAACCTGGGAAAG
GTGATCGACACGCTCACCTGCGGCTTCGCCGACTTGATGGGATACATCCCTCT
GGTGGGGGCCCCTCTGGGCGGAGCCGCGCGCGCCCTGGCTCACGGGGTCCG
GGTGCTCGAGGACGGGGTGAACTACGCCACCGGGAACCTGCCCGGCTGCAGC
TTCTCCATCTTCCTGCTGGCGCTGCTGAGCTGCCTCACCATCCCCGCTAGCGC
ATGA

Figure 3, Codon optimised HCV NS3 polynucleotide

ATGGCCCCCATCACCGCCTACAGCCAGCAGACCCGGGGACTGCTCGGCTGCA
TCATCACCTCTCTGACAGGCCGGGATAAGAACCAGGTGGAGGGCGAGGTGCA
GGTCGTCTCGACCGCTACCCAAAGCTTCCTGGCCACCTGTATCAACGGAGTCT
GCTGGACGGTGTACCATGGCGCCGGCAGCAAGACCCTCGCCGGGCCTAAGGG
CCCCATCACCCAGATGTACACCAACGTGGACCAGGACCTGGTGGGCTGGCAG
GCGCCCCCCGGGGCGAGGAGTATGACCCCATGCACCTGCGGGAGCTCTGACC
TGTATCTGGTGACCAGACATGCCGATGTCATCCCGGTGAGGCGTCGCGGGGAC
AGTAGAGGGAGCCTGCTGAGCCCCCGCCCCGTCAGCTACCTGAAGGGGTCCG
TGGGCGGCCCCCTGCTGTGCCCCTCTGGCCACGTGGTCGGCATCTTCAGGGC
CGCCGTGTGCACGCGCGGCGTGGCCAAGGCCGTGGACTTTATCCCCGTGGAG
AGCATGGAGACCACCATGCGCTCCCCCGTGTTCACCGACAACAGCAGCCCCCC
CGCCGTGCCTCAGACCTTCCAGGTCGCCCACCTCCATGCTCCGACGGGCTCC
GGGAAGTCCACGAAGGTGCCCGCCGCGTACGCGGCCCAGGGATACAAGGTGC
TGGTCCTCAACCCTAGCGTGGCTGCCACACTCGGGTTTGGAGCGTACATGAGC
AAGGCGCACGGCATCGACCCCAACATCAGAACTGGCGTCCGGACCATCACAAC
CGGCGCTCCCATCACTTACTCTACCTACGGCAAGTTCCTGGCTGATGGGGGGT
GTAGTGGGGGCGCGTACGATATTATCATCTGCCAGGAGTGCCACTCTACCGAC
AGCACCACAATCCTGGGCATCGGCACCGTCCTCGACCAGGCTGAGACAGCGG
GCGCCCGCCTGGTGGTGCTGGCCACGGCCACTCCCCCCGGCTCCGTCACGGT
GCCCCACCCCAATATCGAGGAGGTGGCCCTGAGCAACAACGGCGAGATCCCAT
TCTACGGCAAGGCTATCCCGATCGAGGCGATTAAGGGAGGCAGACATCTGATC
TTCTGCCACAGCAAGAAGAAGTGCGACGAGCTCGCCGCCAAGCTGAGCGGCC
TCGGACTCAACGCCGTGGCTTACTACAGGGGACTGGACGTGTCCGTGATCCCG
ACCAGCGGAGACGTGGTGGTCGTGGCCACCGACGCCCTGATGACCGGCTTCA
CCGGAGACTTCGACAGCGTCATCGACTGCAACACCTGCGTGACCCAGACCGTG
GACTTCAGCCTGGACCCCACCTTCACCATCGAGACCACCACAGTGCCCCAGGA
CGCCGTGTCCCGCAGCCAGCGCCGGGGCCGGACCGGCCGCGGCCGGAGTGG
CATCTATAGGTTCGTGACCCCGGGCGAGCGCCCCAGCGGCATGTTCGATAGTT
CCGTGCTGTGCGAGTGCTACGACGCCGGATGCGCGTGGTACGAGCTGACCCC
GGCGGAGACCTCTGTCCGCCTGAGGGCTTACTTGAATACCCCGGGCCTGCCC

GTGTGCCAGGATCATCTCGAGTTCTGGGAATCCGTCTTCACCGGCCTGACACA
CATCGACGCCCATTTCTTGTCCCAAACCAAGCAGGCTGGCGACAATTTCCCGTA
TCTGGTCGCGTACCAGGCCACGGTGTGCGCGCGTGCGCAGGCTCCCCCCCCT
AGCTGGGATCAGATGTGGAAGTGCCTGATCCGCCTGAAGCCCACCCTGCATGG
GCCCACCCCCCTGCTGTACCGCCTGGGCGCGGTGCAGAACGAAGTCACCTTG
ACCCACCCCATCACCAAGTACATCATGGCGTGCATGTCCGCTGACCTGGAGGT
GGTCACCTGA

Figure 4, codon optimised HCV NS4B polynucleotide

ATGTTTTGGGCCAAGCATATGTGGAACTTCATCAGCGGCATCCAGTACCTCGCC
GGGCTGAGCACCCTCCCGGGCAACCCCGCGATCGCAAGCCTGATGGCGTTCA
CAGCGAGCATCACCTCCCCCCTGACTACCCAGAACACACTGCTGTTCAACATCC
TGGGGGGCTGGGTCGCCGCTCAGCTGGCCCCTCCTTCCGCCGCCAGCGCCTT
TGTGGGGGCGGGAATCGCCGGGGCCGCCGTCGGCTCCATCGGACTGGGCAA
GGTGCTGGTCGACATCCTGGCGGGCTACGGCGCGGGAGTCGCCGGAGCCCT
GGTGGCCTTCAAGGTGATGAGCGGAGAGGTGCCAAGCACTGAGGACCTGGTG
AACCTGCTGCCGGCGATCCTGAGCCCGGGCGCCCTGGTGGTGGGCGTGGTGT
GTGCTGCCATCCTCAGGCGCCACGTGGGCCCGGGCGAGGGAGCCGTGCAGT
GGATGAACCGCCTGATCGCCTTTGCCTCCCGCGGCAACCACGTCAGCCCTACA
CATTACGTGCCCGAGAGCGATGCCGCCGCCCGCGTGACCCAGATCCTGAGCT
CCCTGACCATCACCCAGCTGCTCAAGAGGCTGCACCAGTGGATCAACGAGGAC
TGCTCCACCCCTTGCTGA

Figure 5, codon optimised HCV NS5B polynucleotide

ATGTCCATGTCCTACACCTGGACCGGCGCCCTGATCACCCCCTGCGCCGCCGA
GGAGAGCAAGCTCCCGATTAACCCCCTGTCCAACTCTCTGCTCCGCCATCACA
ACATGGTGTATGCCACCACCTCCCGCTCTGCGAGCCTCCGCCAGAAGAAGGTG
ACGTTCGACAGACTGCAGGTGCTGGACGACCATTACAGGGACGTGCTGAAGGA
AATGAAGGCCAAGGCTAGCACCGTGAAGGCCAAGCTGCTCAGCATTGAGGAGG
CTTGCAAGCTGACCCCCCCCCACAGTGCTAAATCCAAGTTCGGCTACGGCGCC
AAGGACGTGAGGAACCTGTCCTCGCGCGCTGTGAACCATATCCGCAGCGTGTG
GGAGGACCTGCTCGAGGACACCGAGACCCCCATCGACACAACCATCATGGCCA
AGTCCGAGGTGTTCTGCGTGCAGCCGGAGAAAGGAGGCCGCAAGCCAGCCCG
CCTGATCGTCTTCCCCGACCTGGGCGTGAGAGTCTGCGAGAAGATGGCCCTCT
ACGACGTGGTGTCCACCCTGCCGCAGGCCGTGATGGGGAGTTCCTACGGCTT
CCAGTACAGCCCGAAGCAGAGGGTGGAGTTCCTGGTGAACACGTGGAAGTCTA
AGAAATGCCCCATGGGGTTCAGTTACGGAACAAGGTGCTTCGGGAGTACTGTG
ACCGAATCCGATATCCGCGTGGAGGAGAGCATCTACCAGTGTTGTGACCTCGC
CCCCGAGGCGAGACAGGCCATCCGCTCCCTGACCGAGAGGCTGTATATCGGC
GGCCCACTGACCAACAGCAAGGGGCAGAACTGCGGCTATCGCCGTTGTCGGG
CCTCCGGGGTGCTCACCACCTCTTGCGGGAACACCCTCACCTGCTACCTCAAG
GCGACCGCTGCCTGCAGAGCCGCGAAGCTGCAGGACTGCACCATGCTCGTGA
ACGGCGACGATCTGGTGGTGATCTGTGAGTCCGCGGGCACGCAGGAGGACGC
GGCGGCCCTGCGGGCGTTCACAGAGGCCATGACACGCTACAGTGCCCCCCCC
GGCGACCCCCCCCCAGCCCGAATACGATCTGGAGCTCATCACTAGTTGCAGCTC
GAACGTGTCTGTGGCCCATGACGCTTCTGGCAAACGGGTGTATTATCTGACGC
GCGATCCCACCACCCCCCTCGCCAGAGCCGCGTGGGAGACAGCTCGGCACAC
CCCTGTGAACTCTTGGCTGGGCAACATCATCATGTACGCCCCTACCCTGTGGG
CTCGCATGATCCTGATGACCCACTTCTTCAGTATCCTCCTCGCTCAGGAGCAGC
TGGAGAAGGCGCTCGACTGCCAGATCTACGGCGCCTGCTATAGTATCGAGCCT
CTCGACCTGCCCCAGATCATCGAGAGACTGCATGGGCTCAGCGCCTTCTCCCT
CCATAGTTACTCTCCTGGAGAAATTAACCGGGTGGCGAGCTGTCTGCGGAAGC
TCGGCGTCCCCCCTCTGCGCGTTTGGCGGCATCGCGCCAGGAGTGTGAGGGC
CAAGCTGCTGAGCCAGGGCGGAAGGGCCGCCACCTGCGGCCGGTATCTCTTC
AACTGGGCCGTGCGCACCAAGCTCAAGCTCACCCCCATCCCTGCCGCCAGTCA
GCTGGATCTCAGTGGGTGGTTCGTGGCCGGCTATTCTGGCGGCGACATCTACC
ACTCCCTCAGCAGGGCGCGCCCCCGCTGGTTCCCCCTGTGCCTGCTGCTCCT
GAGCGTCGGAGTCGGCATCTACCTGCTGCCCAACCGCTGA

Figure 6, *Translation of HCV J4L6 genome (wild-type sequence)*

```
   1 MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR KASERSQPRG
  61 RRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP RGSRPSWGPT DPRRRSRNLG
 121 KVIDTLTCGF ADLMGYIPLV GAPLGGAARA LAHGVRVLED GVNYATGNLP GCSFSIFLLA
 181 LLSCLTIPAS AYEVRNVSGI YHVTNDCSNS SIVYEAADVI MHTPGCVPCV QEGNSSRCWV
 241 ALTPTLAARN ASVPTTTIRR HVDLLVGTAA FCSAMYVGDL CGSIFLVSQL FTFSPRRHET
 301 VQDCNCSIYP GHVSGHRMAW DMMMNWSPTT ALVVSQLLRI PQAVVDMVAG AHWGVLAGLA
 361 YYSMVGNWAK VLIVALLFAG VDGETHTTGR VAGHTTSGFT SLFSSGASQK IQLVNTNGSW
 421 HINRTALNCN DSLQTGFFAA LFYAHKFNSS GCPERMASCR PIDWFAQGWG PITYTKPNSS
 481 DQRPYCWHYA PRPCGVVPAS QVCGPVYCFT PSPVVVGTTD RSGVPTYSWG ENETDVMLLN
 541 NTRPPQGNWF GCTWMNSTGF TKTCGGPPCN IGGVGNRTLI CPTDCFRKHP EATYTKCGSG
 601 PWLTPRCLVD YPYRLWHYPC TLNFSIFKVR MYVGGVEHRL NAACNWTRGE RCNLEDRDRS
 661 ELSPLLLSTT EWQILPCAFT TLPALSTGLI HLHQNIVDVQ YLYGVGSAFV SFAIKWEYIL
 721 LLFLLLADAR VCACLWMMLL IAQAEAALEN LVVLNAASVA GAHGILSFLV FFCAAWYIKG
 781 RLAPGAAYAF YGVWPLLLLL LALPPRAYAL DREMAASCGG AVLVGLVFLT LSPYYKVFLT
 841 RLIWWLQYFI TRAEAHMQVW VPPLNVRGGR DAIILLTCAV HPELIFDITK LLLAILGPLM
 901 VLQAGITRVP YFVRAQGLIR ACMLVRKVAG GHYVQMVFMK LGALTGTYVY NHLTPLRDWA
 961 HAGLRDLAVA VEPVVFSAME TKVITWGADT AACGDIILGL PVSARRGKEI FLGPADSLEG
1021 QGWRLLAPIT AYSQQTRGVL GCIITSLTGR DKNQVEGEVQ VVSTATQSFL ATCINGVCWT
1081 VYHGAGSKTL AGPKGPITQM YTNVDLDLVG WQAPPGARSM TPCSCGSSDL YLVTRHADVI
1141 PVRRRGDSRG SLLSPRPVSY LKGSSGGPLL CPSGHVVGVF RAAVCTRGVA KAVDFIPVES
1201 METTMRSPVF TDNSTPPAVP QTFQVAHLHA PTGSGKSTKV PAAYAAQGYK VLVLNPSVAA
1261 TLGFGAYMSK AHGIDPNIRT GVRTITTGGS ITYSTYGKFL ADGGCSGGAY DIIICDECHS
1321 TDSTTILGIG TVLDQAETAG ARLVVLATAT PPGSVTVPHP NIEEIGLSNN GEIPFYGKAI
1381 PIEAIKGGRH LIFCHSKKKC DELAAKLTGL GLNAVAYYRG LDVSVIPPIG DVVVVATDAL
1441 MTGFTGDFDS VIDCNTCVTQ TVDFSLDPTF TIETTTVPQD AVSRSQRRGR TGRGRSGIYR
1501 FVTPGERPSG MFDSSVLCEC YDAGCAWYEL TPAETSVRLR AYLNTPGLPV CQDHLEFWES
1561 VFTGLTHIDA HFLSQTKQAG DNFPYLVAYQ ATVCARAQAP PPSWDQMWKC LIRLKPTLHG
1621 PTPLLYRLGA VQNEVILTHP ITKYIMACMS ADLEVVTSTW VLVGGVLAAL AAYCLTTGSV
1681 VIVGRIILSG KPAVVPDREV LYQEFDEMEE CASQLPYIEQ GMQLAEQFKQ KALGLLQTAT
1741 KQAEAAAPVV ESKWRALETF WAKHMWNFIS GIQYLAGLST LPGNPAIASL MAFTASITSP
1801 LTTQNTLLFN ILGGWVAAQL APPSAASAFV GAGIAGAAVG SIGLGKVLVD ILAGYGAGVA
1861 GALVAFKVMS GEVPSTEDLV NLLPAILSPG ALVVGVVCAA ILRRHVGPGE GAVQWMNRLI
1921 AFASRGNHVS PTHYVPESDA AARVTQILSS LTITQLLKRL HQWINEDCST PCSGSWLRDV
1981 WDWICTVLTD FKTWLQSKLL PRLPGVPFLS CQRGYKGVWR GDGIMQTTCP CGAQIAGHVK
2041 NGSMRIVGPR TCSNTWHGTF PINAYTTGPC TPSPAPNYSR ALWRVAAEEY VEVTRVGDFH
2101 YVTGMTTDNV KCPCQVPAPE FFTEVDGVRL HRYAPACKPL LREDVTFQVG LNQYLVGSQL
```

2161 PCEPEPDVTV LTSMLTDPSH ITAETAKRRL ARGSPPSLAS SSASQLSAPS LKATCTTHHD
2221 SPDADLIEAN LLWRQEMGGN ITRVESENKV VILDSFEPLH AEGDEREISV AAEILRKSRK
2281 FPSALPIWAR PDYNPPLLES WKDPDYVPPV VHGCPLPPTK APPIPPPRRK RTVVLTESNV
2341 SSALAELATK TFGSSGSSAV DSGTATALPD LASDDGDKGS DVESYSSMPP LEGEPGDPDL
2401 SDGSWSTVSE EASEDVVCCS MSYTWTGALI TPCAAEESKL PINPLSNSLL RHHNMVYATT
2461 SRSASLRQKK VTFDRLQVLD DHYRDVLKEM KAKASTVKAK LLSIEEACKL TPPHSAKSKF
2521 GYGAKDVRNL SSRAVNHIRS VWEDLLEDTE TPIDTTIMAK SEVFCVQPEK GGRKPARLIV
2581 FPDLGVRVCE KMALYDVVST LPQAVMGSSY GFQYSPKQRV EFLVNTWKSK KCPMGFSYDT
2641 RCFDSTVTES DIRVEESIYQ CCDLAPEARQ AIRSLTERLY IGGPLTNSKG QNCGYRRCRA
2701 SGVLTTSCGN TLTCYLKATA ACRAAKLQDC TMLVNGDDLV VICESAGTQE DAAALRAFTE
2761 AMTRYSAPPG DPPQPEYDLE LITSCSSNVS VAHDASGKRV YYLTRDPTTP LARAAWETAR
2821 HTPINSWLGN IIMYAPTLWA RMILMTHFFS ILLAQEQLEK ALDCQIYGAC YSIEPLDLPQ
2881 IIERLHGLSA FTLHSYSPGE INRVASCLRK LGVPPLRTWR HRARSVRAKL LSQGGRAATC
2941 GRYLFNWAVR TKLKLTPIPA ASQLDLSGWF VAGYSGGDIY HSLSRARPRW FPLCLLLLSV
3001 GVGIYLLPNR

PG5029

Figure 7, p7313-ie

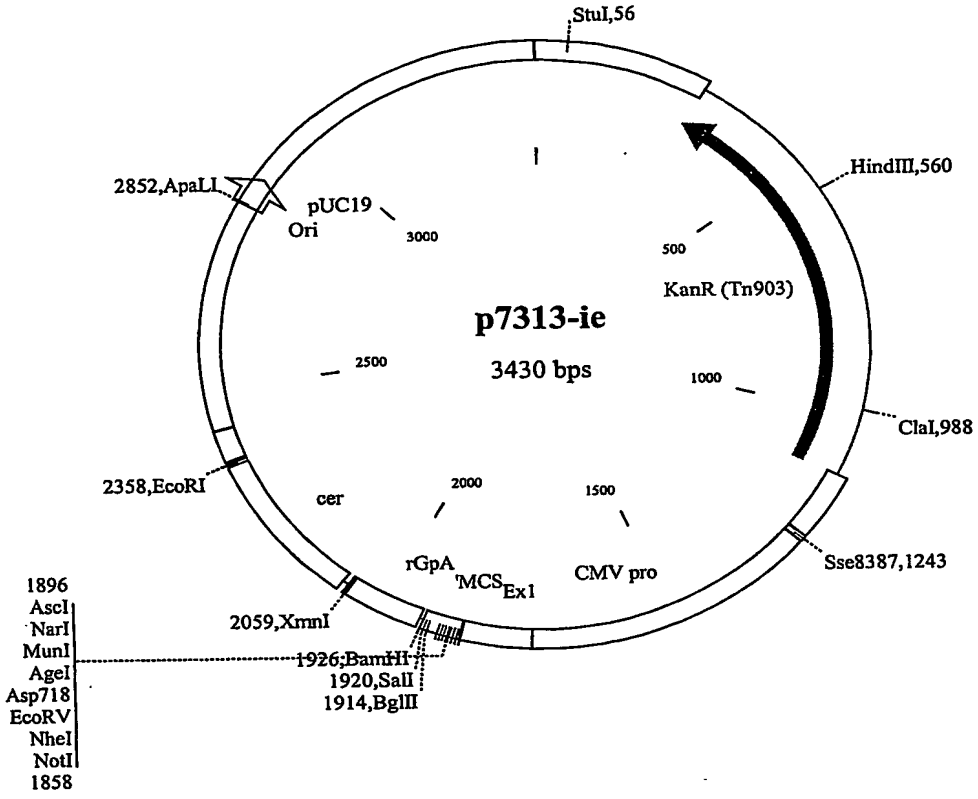
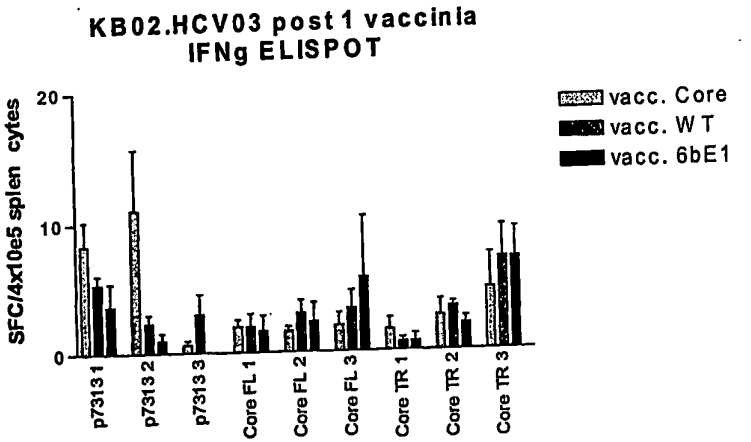
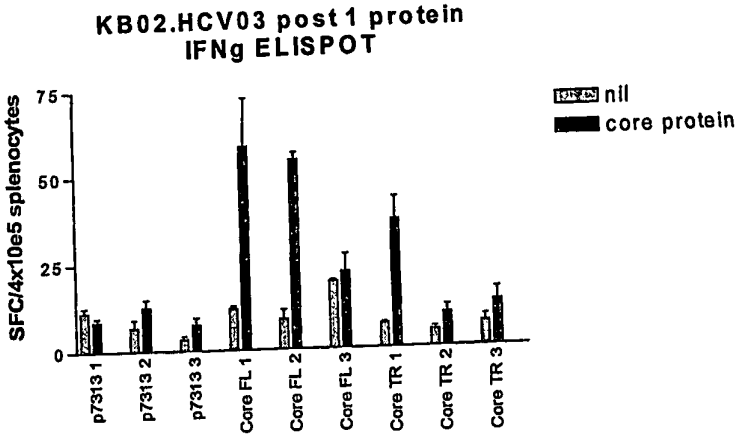
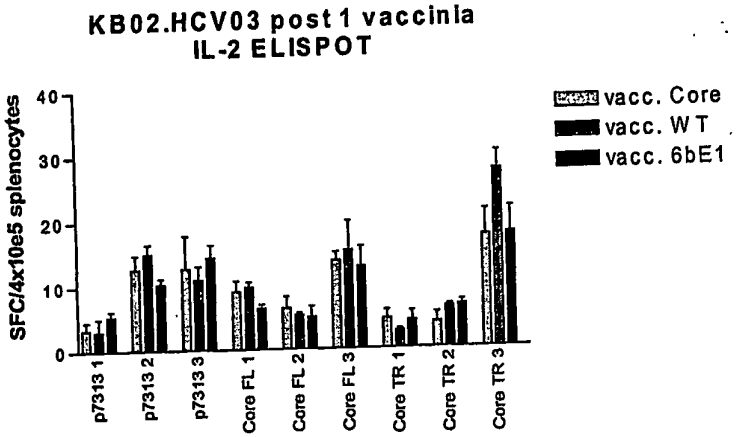


Figure 8, Immune responses to Core







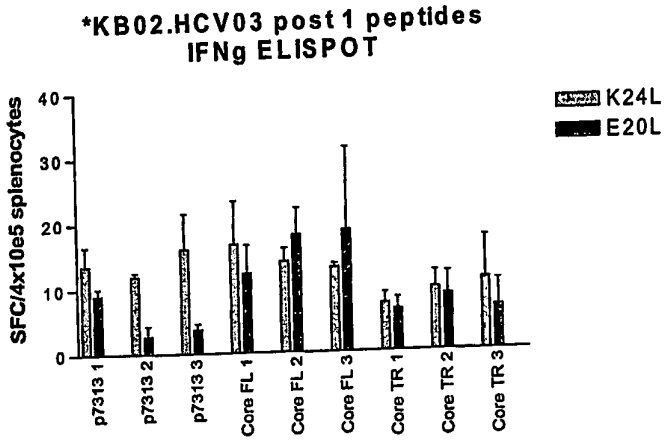
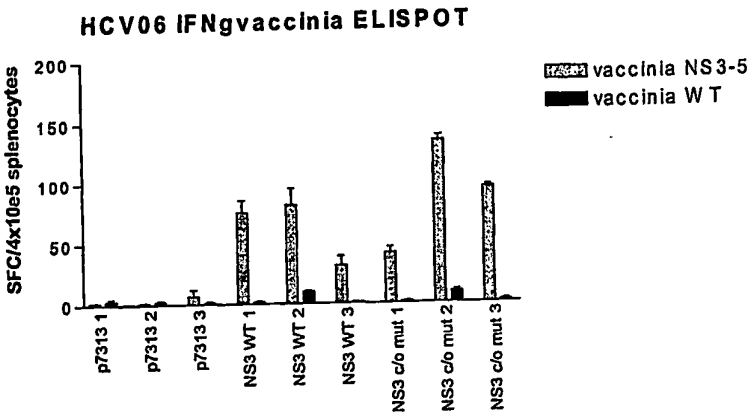
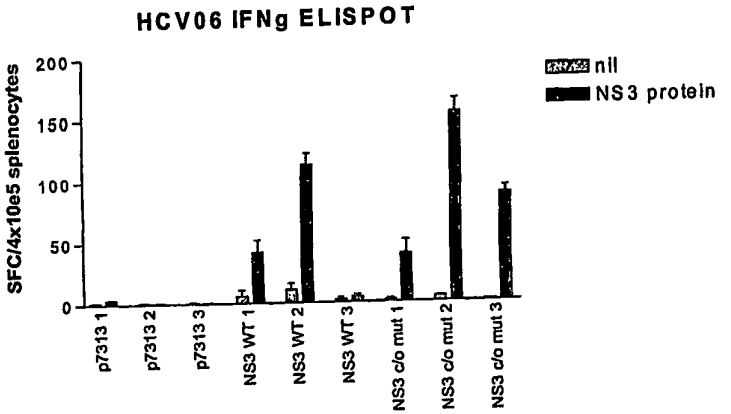
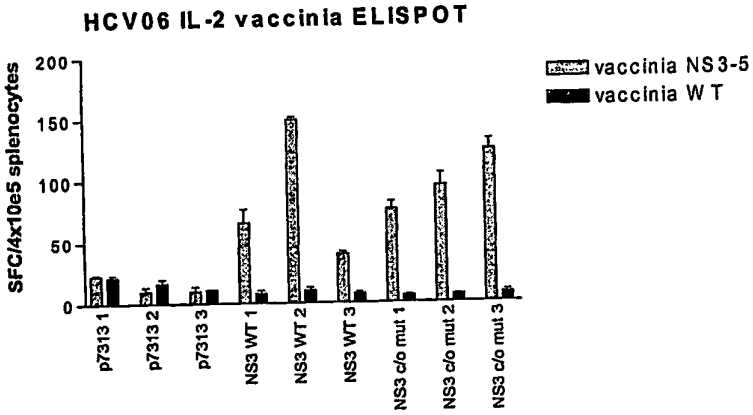


Figure 9, NS3 immunogenicity







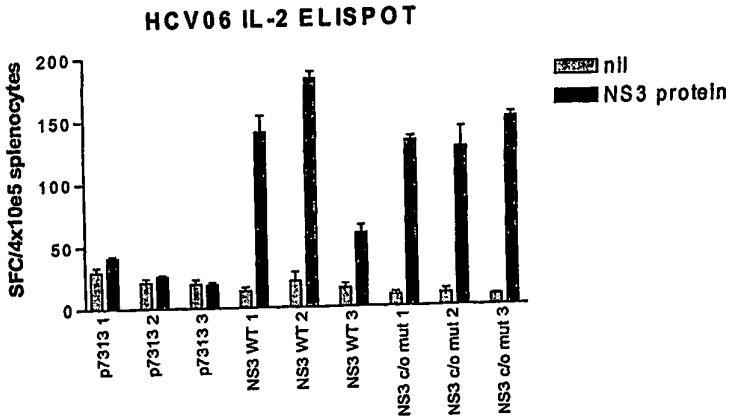
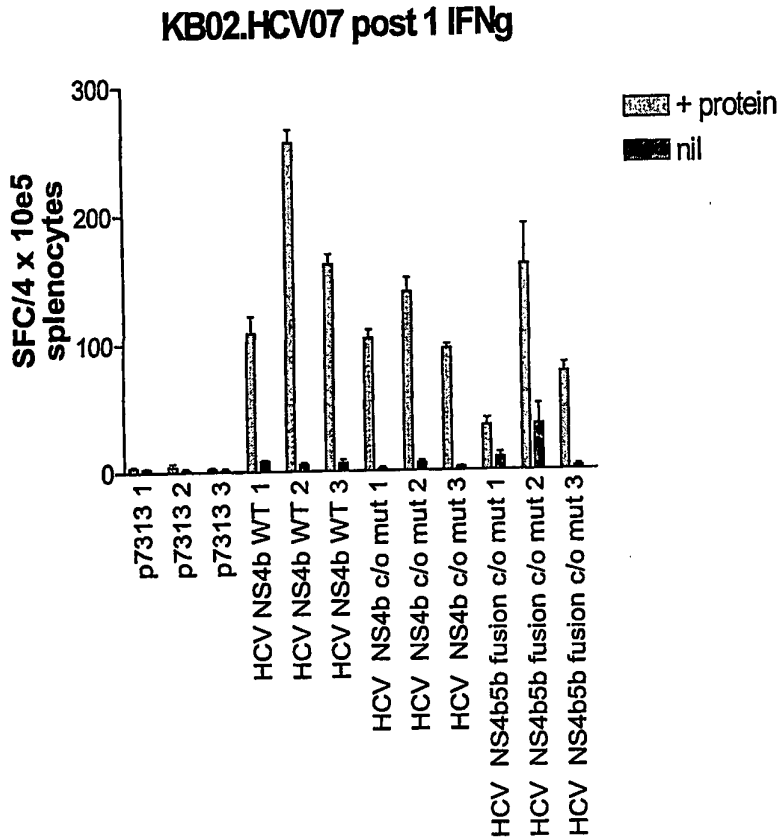


Figure 10, Immune responses to NS4B



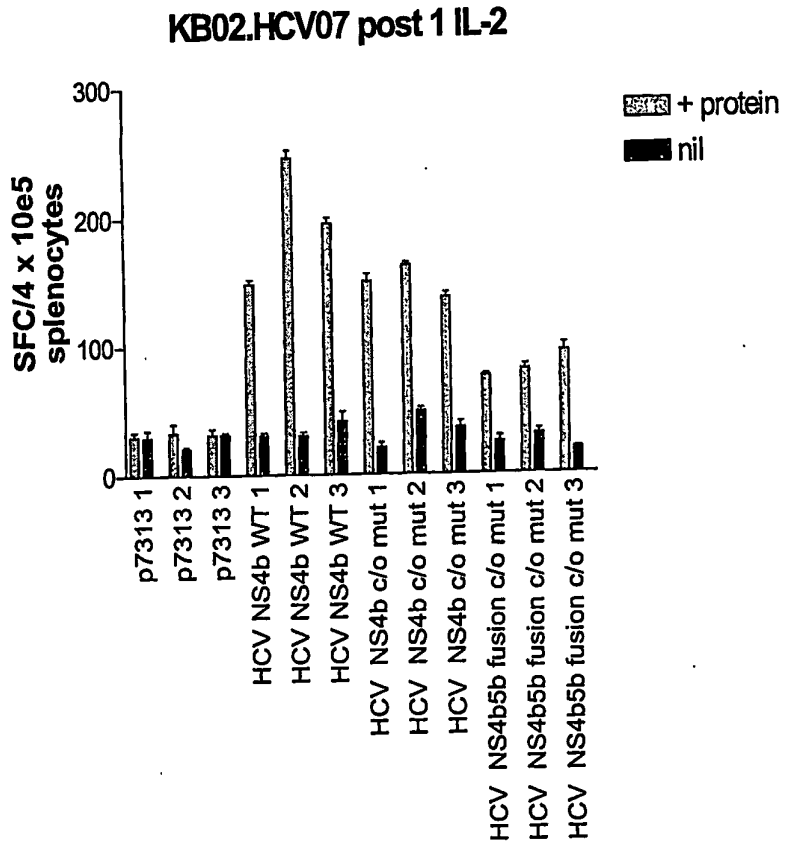
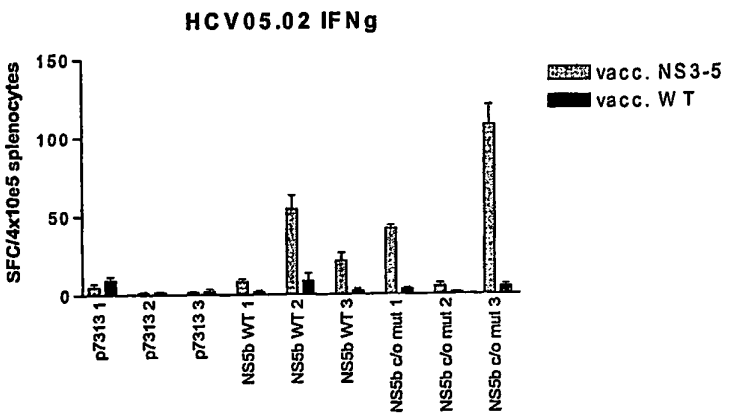
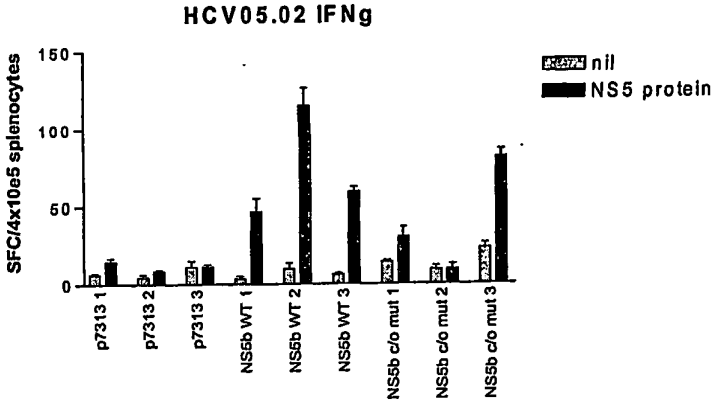
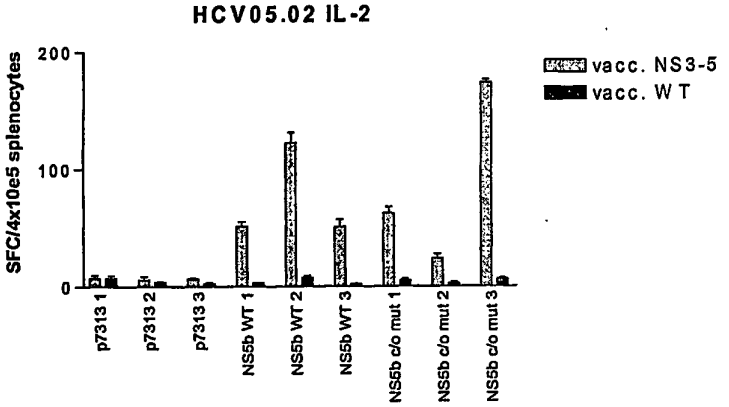
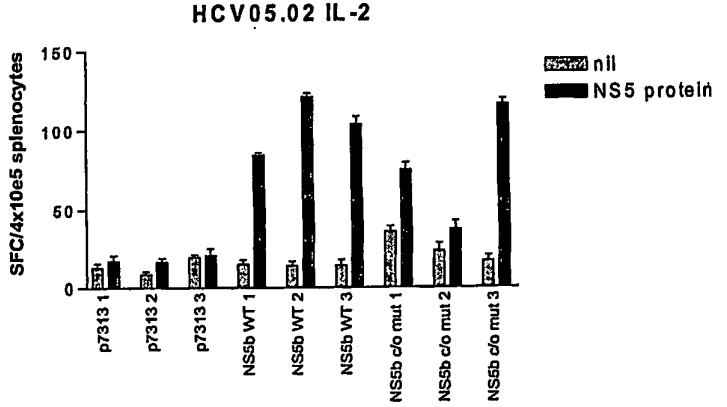


Figure 11, NS5B immune responses









BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- [x] BLACK BORDERS
- [x] IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- [x] FADED TEXT OR DRAWING
- [ ] BLURED OR ILLEGIBLE TEXT OR DRAWING
- [ ] SKEWED/SLANTED IMAGES
- [x] COLORED OR BLACK AND WHITE PHOTOGRAPHS
- [ ] GRAY SCALE DOCUMENTS
- [ ] LINES OR MARKS ON ORIGINAL DOCUMENT
- [ ] REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- [ ] OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox